

Distributions with fixed marginals maximizing the mass of the endograph of a function

Thomas Mroz^a, Wolfgang Trutschnig^{a,*}, Juan Fernández Sánchez^b

^a*Department of Mathematics, University of Salzburg, Hellbrunner Strasse 34, 5020 Salzburg, Austria,
Tel.: +43 662 8044-5312, Fax: +43 662 8044-137*

^b*Grupo de Investigación de Análisis Matemático, Universidad de Almería, La Cañada de San Urbano,
Almería, Spain*

Abstract

We solve the problem of maximizing the probability that X does not default before Y within the class of all random variables X, Y with given distribution functions F and G respectively, and construct a dependence structure attaining the maximum. After translating the maximization problem to the copula setting we generalize it and prove that for each (not necessarily monotonic) transformation $T : [0, 1] \rightarrow [0, 1]$ there exists a completely dependent copula maximizing the mass of the endograph $\Gamma^{\leq}(T)$ of T and derive a simple and easily calculable formula for the maximum. Analogous expressions for the minimal mass are given. Several examples and graphics illustrate the main results and falsify some natural conjectures.

Keywords: Copula, Dependence, Coupling, Endograph, Markov Kernel

2010 MSC: 60E05, 28A50, 91G70

1. Introduction

Suppose that F and G are (continuous) distribution functions of two default times. It is well known from coupling theory (see [16]) that there exists a maximal coupling, i.e. a two-dimensional distribution function H with marginals F and G such that for the case of $(X, Y) \sim H$ the probability of a joint default $\mathbb{P}(X = Y)$ is maximal (within the class of all two-dimensional distribution functions having F and G as marginals). Translating to the class of copulas (see [11] and Section 3), maximizing the probability of a joint default means calculating $\sup_{A \in \mathcal{C}} \mu_A(\Gamma(T))$ for $T : [0, 1] \rightarrow [0, 1]$ being defined by $T = G \circ F^{-}$, F^{-} denoting the quasi-inverse of F , $\Gamma(T)$ the graph of T , \mathcal{C} the family of all two-dimensional copulas and μ_A being the doubly stochastic measure corresponding to the copula $A \in \mathcal{C}$. As pointed out in [11] there is a (not necessarily unique) copula A_0 with

$$\mu_{A_0}(\Gamma(T)) = \sup_{A \in \mathcal{C}} \mu_A(\Gamma(T)) \quad (1)$$

*Corresponding author

Email addresses: thomasmroz@a1.net (Thomas Mroz), wolfgang@trutschnig.net (Wolfgang Trutschnig), juanfernandez@ual.es (Juan Fernández Sánchez)

that can even be computed in closed form. Considering $(U, V) \sim A_0$ and setting $(X, Y) = (F^- \circ U, G^- \circ V)$, the pair (X, Y) has marginal distribution functions F and G and maximizes the joint default probability.

In the current paper we tackle the closely related problem of maximizing $\mathbb{P}(Y \leq X)$, the probability that X does not default before Y , and solve it in a definitive manner. We first translate the maximization problem to the copula setting and prove the existence of a (mutually) completely dependent copula $A_R \in \mathcal{C}$ with

$$\mu_{A_R}(\Gamma^{\leq}(T)) = \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)), \quad (2)$$

where $T = G \circ F^-$ and $\Gamma^{\leq}(T) = \{(x, y) \in [0, 1]^2 : y \leq T(x)\}$ denotes the endograph of T . Afterwards we study the situation of not necessarily monotonic $T : [0, 1] \rightarrow [0, 1]$ and prove a simple and easily calculable formula for $\sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T))$ only involving the distribution function of T . As in the monotonic case it is possible to construct a completely dependent copula maximizing the mass of $\Gamma^{\leq}(T)$. Finally, using the just mentioned results we derive an equally simple formula for $\inf_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T))$ and show that there are situations where the infimum is not attained.

The rest of the paper is organized as follows: Section 2 gathers some preliminaries and notations. In Section 3 we prove the Markov kernel version of Sklar's theorem and then apply it to show that $\mathbb{P}(Y \leq S \circ X) = \mu_A(\Gamma^{\leq}(T))$ holds, where $S : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary measurable transformation, (X, Y) has marginals F, G and copula A , and T is defined by $T = G \circ S \circ F^-$. All aforementioned maximization results are collected in Section 4. Section 5 presents an alternative proof of the main result and derives some useful consequences.

2. Notation and Preliminaries

For every d -dimensional random vector \mathbf{X} on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ we will write $\mathbf{X} \sim F$ if \mathbf{X} has distribution function (d.f., for short) F and let $\mu_F = \mathbb{P}^{\mathbf{X}}$ denote the corresponding distribution on the Borel σ -field $\mathcal{B}(\mathbb{R}^d)$ of \mathbb{R}^d . For every univariate distribution function F we will let F^- denote the quasi-inverse of F , i.e. $F^-(q) = \inf\{x \in \mathbb{R} : F(x) \geq q\}$. Note that for every $q \in (0, 1)$ we have $F^-(q) \leq x$ if and only if $q \leq F(x)$, that for $X \sim F$ and F continuous we have $F \circ X \sim \mathcal{U}(0, 1)$ and that the random variable $F^- \circ F \circ X$ coincides with X with probability one. For further properties of F^- we refer, for instance, to [6]. Given univariate distribution functions F and G , we will let $\mathcal{H}_{F,G}$ denote the Fréchet class of F and G , i.e. the family of all two-dimensional distribution functions having F and G as marginals; $\mathcal{P}_{F,G}$ will denote the corresponding class of probability measures on $\mathcal{B}(\mathbb{R}^2)$. $\mathcal{B}([0, 1])$ and $\mathcal{B}([0, 1]^2)$ denote the Borel σ -fields on $[0, 1]$ and $[0, 1]^2$, λ and λ_2 the Lebesgue measure on $\mathcal{B}([0, 1])$ and $\mathcal{B}([0, 1]^2)$ respectively. For every measurable transformation $T : [0, 1] \rightarrow [0, 1]$ the push-forward of λ via T will be denoted by λ^T .

As already mentioned before, \mathcal{C} will denote the family of all two-dimensional *copulas*. For background on copulas we refer to [3, 13]. M and W will denote upper and the lower Fréchet-Hoeffding bound, Π the product copula. d_∞ will denote the uniform distance on \mathcal{C} ; it is well known that (\mathcal{C}, d_∞) is a compact metric space and that d_∞ is metrization of weak convergence in \mathcal{C} . For every $A \in \mathcal{C}$ μ_A will denote the corresponding *doubly stochastic measure* defined by $\mu_A([0, x] \times [0, y]) = A(x, y)$ for all $x, y \in [0, 1]$, $\mathcal{P}_{\mathcal{C}}$ the class of all these doubly stochastic measures.

A *Markov kernel* from \mathbb{R} to $\mathcal{B}(\mathbb{R})$ is a mapping $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ such that $x \mapsto K(x, B)$ is measurable for every fixed $B \in \mathcal{B}(\mathbb{R})$ and $B \mapsto K(x, B)$ is a probability measure for every fixed $x \in \mathbb{R}$. Given real-valued random variables X, Y on $(\Omega, \mathcal{A}, \mathbb{P})$, a Markov kernel $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ is called a *regular conditional distribution of Y given X* if for every $B \in \mathcal{B}(\mathbb{R})$

$$K(X(\omega), B) = \mathbb{E}(\mathbf{1}_B \circ Y | X)(\omega) \quad (3)$$

holds \mathbb{P} -a.s. It is well known that for each pair (X, Y) of real-valued random variables a regular conditional distribution $K(\cdot, \cdot)$ of Y given X exists, that $K(\cdot, \cdot)$ is unique \mathbb{P}^X -a.s. (i.e. unique for \mathbb{P}^X -almost every $x \in \mathbb{R}$) and that $K(\cdot, \cdot)$ only depends on the distribution $\mathbb{P}^{(X, Y)}$. Hence, given $(X, Y) \sim H$, we will denote (a version of) the regular conditional distribution of Y given X by $K_H(\cdot, \cdot)$ and refer to $K_H(\cdot, \cdot)$ simply as *Markov kernel of H* or *Markov kernel of (X, Y)* . Note that for every two-dimensional distribution function H , its Markov kernel $K_H(\cdot, \cdot)$, and every Borel set $G \in \mathcal{B}(\mathbb{R}^2)$ the following *disintegration* formula holds ($G_x = \{y \in \mathbb{R} : (x, y) \in G\}$ denoting the x -section of G for every $x \in \mathbb{R}$)

$$\int_{\mathbb{R}} K_H(x, G_x) d\lambda(x) = \mu_H(G). \quad (4)$$

For $A \in \mathcal{C}$ we will directly consider the corresponding Markov kernel $K_A(\cdot, \cdot)$ to be defined on $[0, 1] \times \mathcal{B}([0, 1])$. Considering that in this case eq. (4) implies that

$$\int_{[0, 1]} K_A(x, F) d\lambda(x) = \lambda(F) \quad (5)$$

holds for every $F \in \mathcal{B}([0, 1])$, and that, additionally, every Markov kernel $K : [0, 1] \times \mathcal{B}([0, 1]) \rightarrow [0, 1]$ fulfilling eq. (5) obviously induces a unique element $\mu \in \mathcal{P}_{\mathcal{C}}$, it follows that there is a one-to-one correspondence between \mathcal{C} and the family of all Markov kernels $K : [0, 1] \times \mathcal{B}([0, 1]) \rightarrow [0, 1]$ fulfilling eq. (5). Notice that for $A \in \mathcal{C}$ eq. (5) also implies that $K_A(x, \{0, 1\}) = 0$ holds for λ -almost every $x \in [0, 1]$, so it is always possible to choose a (version of the) kernel fulfilling $K_A(x, \{0, 1\}) = 0$ for every $x \in [0, 1]$. For more details and properties of conditional expectation, regular conditional distributions, and disintegration see [8] and [9], various results underlining the usefulness of the Markov kernel perspective can be found in [3] and the references therein.

In the sequel \mathcal{T} will denote the class of all λ -preserving transformations $h : [0, 1] \rightarrow [0, 1]$, \mathcal{T}_b the subset of all bijective $h \in \mathcal{T}$, and \mathcal{T}_l the subset of all piecewise linear, bijective $h \in \mathcal{T}$. A copula $A \in \mathcal{C}$ will be called *completely dependent* if and only if there exists $h \in \mathcal{T}$ such that $K(x, E) = \mathbf{1}_E(h(x))$ is a regular conditional distribution of A (see [10, 17] for equivalent definitions and main properties). For every $h \in \mathcal{T}$ the induced completely dependent copula will be denoted by A_h throughout the rest of the paper, \mathcal{C}_d will denote the family of all completely dependent copulas.

Following [3, 18], for every $h \in \mathcal{T}$ and every copula $A \in \mathcal{C}$ we will let $\mathcal{S}_h(A) \in \mathcal{C}$ denote the (generalized) *h -shuffle* of A , defined implicitly via the corresponding doubly stochastic measures by

$$\mu_{\mathcal{S}_h(A)}(E \times F) = \mu_A(h^{-1}(E) \times F) \quad (6)$$

for all $E, F \in \mathcal{B}([0, 1])$. Notice that $\mathcal{S}_h(A)$ is a shuffle in the sense of [2] if $h \in \mathcal{T}_b$, and that it is a shuffle in the sense of [12] (to which we will refer as classical shuffle in the sequel) if $h \in \mathcal{T}_l$.

3. Markov kernel version of Sklar's theorem

Suppose now that the vector (X, Y) has distribution function $H \in \mathcal{H}_{F,G}$ with F, G continuous. According to Sklar's theorem (see [3, 13]) there exists a unique copula $A \in \mathcal{C}$ such that $H(x, y) = A(F(x), G(y))$ holds for all $x, y \in \mathbb{R}$. Translating this to the Markov kernel setting we get the following result describing how to construct a kernel of H given the kernel of A :

Lemma 1. *Suppose that F, G are continuous distribution functions, that (X, Y) has d.f. $H \in \mathcal{H}_{F,G}$ and copula A , and let $K_A(\cdot, \cdot)$ denote a Markov kernel of A fulfilling $K_A(x, \{0, 1\}) = 0$ for all $x \in [0, 1]$. Then setting*

$$K(x, (-\infty, y]) := K_A(F(x), [0, G(y)]) \quad (7)$$

for all $x, y \in \mathbb{R}$ defines a Markov kernel $K(\cdot, \cdot)$ of $(X, Y) \sim H$.

Proof: (i) We need to show that $y \mapsto K(x, (-\infty, y])$ is a distribution function for every fixed $x \in \mathbb{R}$: The fact that $y \mapsto K(x, (-\infty, y])$ is non-decreasing is trivial. If $(y_n)_{n \in \mathbb{N}}$ is monotonically decreasing with limit $y \in \mathbb{R}$ then, using continuity of G and the fact that $K_A(\cdot, \cdot)$ is a Markov kernel, we have $\bigcap_{n=1}^{\infty} [0, G(y_n)] = [0, G(y)]$ as well as

$$\lim_{n \rightarrow \infty} K(x, (-\infty, y_n]) = \lim_{n \rightarrow \infty} K_A(F(x), [0, G(y_n)]) = K_A(F(x), [0, G(y)]) = K(x, (-\infty, y])$$

Since both $\lim_{y \rightarrow -\infty} K(x, (-\infty, y]) = 0$ and $\lim_{y \rightarrow \infty} K(x, (-\infty, y]) = 1$ follow in the same manner, $y \mapsto K(x, (-\infty, y])$ is a distribution function and we can extend $K(x, \cdot)$ from the generator $\mathcal{E} = \{(-\infty, y] : y \in \mathbb{R}\}$ to a probability measure on $\mathcal{B}(\mathbb{R})$ in the standard way ([9]).

(ii) Measurability of $x \mapsto K(x, (-\infty, y])$ for every fixed $y \in \mathbb{R}$ is a direct consequence of measurability of $x \mapsto F(x)$ and the fact that $K_A(\cdot, \cdot)$ is a Markov kernel. Considering that $\mathcal{D} = \{E \subseteq \mathbb{R} : x \mapsto K(x, E) \text{ measurable}\}$ is a Dynkin system, that \mathcal{E} is closed w.r.t. intersection, and that $\mathcal{E} \subseteq \mathcal{D}$, it follows that $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_\sigma(\mathcal{E}) \subseteq \mathcal{D}$ (see [9]), implying that $K(\cdot, \cdot)$ is indeed a Markov kernel.

(iii) It remains to show that $K(\cdot, \cdot)$ is a Markov kernel of (X, Y) . Let $q \in (0, 1)$ be arbitrary but fixed. Then setting $I = \int_{(-\infty, F^-(q)]} K(t, (-\infty, y]) d\mathbb{P}^X(t)$ and using disintegration, continuity of F , change of coordinates and Sklar's theorem we get

$$\begin{aligned} I &= \int_{\mathbb{R}} \mathbf{1}_{(-\infty, F^-(q))}(t) K_A(F(t), [0, G(y)]) d\mathbb{P}^X(t) \\ &= \int_{\mathbb{R}} \mathbf{1}_{(-\infty, q)}(F(t)) K_A(F(t), [0, G(y)]) d\mathbb{P}^X(t) \\ &= \int_{[0, 1]} \mathbf{1}_{(-\infty, q)}(z) K_A(z, [0, G(y)]) d\mathbb{P}^{F \circ X}(z) \\ &= \int_{[0, 1]} \mathbf{1}_{[0, q)}(z) K_A(z, [0, G(y)]) d\lambda(z) = \int_{[0, q]} K_A(z, [0, G(y)]) d\lambda(z) \\ &= A(q, G(y)) = A(F \circ F^-(q), G(y)) = H(F^-(q), y). \end{aligned}$$

This shows that we have

$$H(z, y) = \int_{(-\infty, z]} K(t, (-\infty, y]) d\mathbb{P}^X(t) \quad (8)$$

for all $y \in \mathbb{R}$ and all z of the form $z = F^-(q)$ for some $q \in (0, 1)$. In case of $q = 1$ and $F^-(1) < \infty$ we can use completely the same arguments to show that

$$\int_{(-\infty, F^-(1)]} K(t, (-\infty, y]) d\mathbb{P}^X(t) = H(F^-(1), y) = G(y) \quad (9)$$

holds for every $y \in \mathbb{R}$. The extension to full \mathbb{R}^2 is now straightforward: Let $x, y \in \mathbb{R}$ be arbitrary, set $q := F(x)$ and $z := F^-(q)$. If $q \in (0, 1)$ then $z \leq x$ as well as $\mathbb{P}(X \in (z, x]) = 0$ follow and we get

$$\begin{aligned} H(x, y) &= \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq z, Y \leq y) = \int_{(-\infty, z]} K(t, (-\infty, y]) d\mathbb{P}^X(t) \\ &= \int_{(-\infty, x]} K(t, (-\infty, y]) d\mathbb{P}^X(t). \end{aligned}$$

In case of $q = 1$ we have $F^-(1) \leq x < \infty$, so using eq. (9) and $F \circ F^-(1) = 1$ we get

$$H(x, y) = G(y) = \int_{(-\infty, F^-(1)]} K(t, (-\infty, y]) d\mathbb{P}^X(t) = \int_{(-\infty, x]} K(t, (-\infty, y]) d\mathbb{P}^X(t),$$

and in case of $q = 0$ it follows that

$$H(x, y) = 0 = \int_{(-\infty, x]} K(t, (-\infty, y]) d\mathbb{P}^X(t).$$

Altogether we have shown that $H(x, y) = \int_{(-\infty, x]} K(t, (-\infty, y]) d\mathbb{P}^X(t)$ holds for all $x, y \in \mathbb{R}$, so, extending in the standard way from \mathcal{E}^2 to $\mathcal{B}(\mathbb{R}^2)$ (see [8]) we get that $K(\cdot, \cdot)$ is a Markov kernel of (X, Y) . ■

Proceeding analogously to the proof of Lemma 1 we can show the following result, describing how to construct a kernel $K_A(\cdot, \cdot)$ of the copula A if the kernel $K_H(\cdot, \cdot)$ of (X, Y) is known:

Lemma 2. *Suppose that F, G are continuous distribution functions, that (X, Y) has d.f. $H \in \mathcal{H}_{F,G}$ and copula A , and let $K_H(\cdot, \cdot)$ denote a Markov kernel of H with $K_H(x, (G^-(0), G^-(1))) = 1$ for every $x \in \mathbb{R}$. Then setting*

$$K(x, [0, y)) := K_H(F^-(x), (-\infty, G^-(y))) \quad (10)$$

for all $x \in (0, 1)$ and $y \in [0, 1]$ defines a Markov kernel $K(\cdot, \cdot)$ of $A \in \mathcal{C}$.

Suppose now that $S : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary Borel-measurable mapping. In the sequel we will let $\Gamma(S)$ and $\Gamma^{\leq}(S)$ denote the graph and the endograph of S respectively, i.e.

$$\Gamma(S) = \{(x, S(x)) : x \in \mathbb{R}\}, \quad \Gamma^{\leq}(S) = \{(x, y) \in \mathbb{R}^2 : y \leq S(x)\}. \quad (11)$$

Lemma 1 allows to express $\mathbb{P}(Y \leq X)$ as well as $\mathbb{P}(Y = X)$ in terms of F, G and the underlying copula A . In order to prove a more general result and to simplify notation, given (continuous) F, G and (measurable) S we will write

$$T := G \circ S \circ F^- \quad (12)$$

in the sequel. In general T is only well-defined on $(0, 1)$ - we will however, directly consider it as function on $[0, 1]$ by setting $T(0) = 0$ and $T(1) = 1$.

Theorem 3. Suppose that X, Y are random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ with joint distribution function H , continuous marginals F and G and copula A . Furthermore let $S : \mathbb{R} \rightarrow \mathbb{R}$ be an arbitrary Borel-measurable mapping and define T according to eq. (12). Then the following identities hold:

$$\mathbb{P}^{(X,Y)}(\Gamma(S)) = \mu_A(\Gamma(T)), \quad \mathbb{P}^{(X,Y)}(\Gamma^{\leq}(S)) = \mu_A(\Gamma^{\leq}(T)) \quad (13)$$

Proof: Using the fact that $\mathbb{P}(F^- \circ F \circ X = X) = 1$, disintegration and Lemma 1 the second identity can be proved as follows:

$$\begin{aligned} \mathbb{P}^{(X,Y)}(\Gamma^{\leq}(S)) &= \int_{\Omega} K(X(\omega), (-\infty, S \circ X(\omega)]) d\mathbb{P}(\omega) \\ &= \int_{\Omega} K_A(F \circ X(\omega), [0, G \circ S \circ F^- \circ F \circ X(\omega)]) d\mathbb{P}(\omega) \\ &= \int_{\mathbb{R}} K_A(z, [0, G \circ S \circ F^-(z)]) d\mathbb{P}^{F \circ X}(z) \\ &= \int_{[0,1]} K_A(z, [0, T(z)]) d\lambda(z) = \mu_A(\Gamma^{\leq}(T)). \end{aligned}$$

Working with $K(X(\omega), \{S \circ X(\omega)\})$ instead of $K(X(\omega), (-\infty, S \circ X(\omega)])$ the first identity $\mathbb{P}^{(X,Y)}(\Gamma(S)) = \mu_A(\Gamma(T))$ follows in the same manner. ■

4. Maximizing the mass of the endograph

For the special case of $S = id_{\mathbb{R}}$ calculating $\sup_{\mu \in \mathcal{P}(F,G)} \mu(\Gamma^{\leq}(S))$ corresponds to finding (joint) distributions of (X, Y) for which $\mathbb{P}(Y \leq X)$ is as big as possible - interpreting X, Y as lifetimes or default times of financial institutions this translates to maximizing the probability that X does not die or default before Y . Notice that, setting $\psi(x, y) = x + y$ and considering the pair $(-X, Y)$, this maximization problem can be considered a special case of the much more general situation studied in [4, 5]. Theorem 3 implies that the problem can be reduced to the family of copulas, i.e. we have

$$\overline{m} := \sup_{\mu \in \mathcal{P}(F,G)} \mu(\Gamma^{\leq}(id_{\mathbb{R}})) = \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)). \quad (14)$$

Obviously the same is true when considering minimal probabilities, i.e.

$$\underline{m} := \inf_{\mu \in \mathcal{P}(F,G)} \mu(\Gamma^{\leq}(id_{\mathbb{R}})) = \inf_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)) \quad (15)$$

holds. Taking into account that in case of $S = id_{\mathbb{R}}$ the mapping $T = G \circ S \circ F^-$ according to Theorem 3 is non-decreasing, it is actually possible to calculate \overline{m} and even construct a dependence structure for which $\mathbb{P}(Y \leq X)$ coincides with \overline{m} . The following result holds:

Theorem 4. Suppose that $T : [0, 1] \rightarrow [0, 1]$ is non-decreasing. Then we have

$$\sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)) = 1 + \inf_{x \in [0,1]} (T(x) - x). \quad (16)$$

Additionally, setting $z = \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T))$ and letting $R \in \mathcal{T}$ denote the rotation $R(x) = x + z \pmod{1}$, we have $\mu_{AR}(\Gamma^{\leq}(T)) = \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T))$.

Proof: Considering $\Gamma^\leq(T) \subseteq [0, x] \times [0, T(x)] \cup [x, 1] \times [0, 1]$ it follows that $\mu_A(\Gamma^\leq(T)) \leq T(x) + 1 - x$ holds for every $x \in [0, 1]$ and every $A \in \mathcal{C}$, which implies that the left-hand side of (16) is smaller than or equal to the right-hand side.

To prove the other inequality set $z = \inf_{x \in [0, 1]} (T(x) + 1 - x)$. In case of $z = 1$ we have $T(x) \geq x$ for every x so considering $\mu_M(\Gamma^\leq(T)) = 1$ we are done. Suppose now that $z < 1$. Compactness of $[0, 1]$ implies the existence of a sequence $(x_n)_{n \in \mathbb{N}}$ and a point $x^* \in [0, 1]$ such that $\lim_{n \rightarrow \infty} x_n = x^*$ and $\lim_{n \rightarrow \infty} (T(x_n) + 1 - x_n) = z$. Using $z < 1$ we get $x^* > 0$ and, using monotonicity of T it follows that $T(x^* -) + 1 - x^* = z$. Letting $R : [0, 1] \rightarrow [0, 1]$ denote the rotation mentioned in the theorem, obviously $R \in \mathcal{T}$ holds. Considering that for every $x \in [x^* - T(x^* -), 1]$ we have

$$\begin{aligned} R(x) &= T(x^* -) - x^* + x = T(x^* -) + 1 - x^* - 1 + x \\ &\leq T(x) + 1 - x - 1 + x = T(x) \end{aligned}$$

it follows immediately that

$$\mu_{A_R}(\Gamma^\leq(T)) \geq 1 - (x^* - T(x^* -)) = z = \inf_{x \in [0, 1]} (T(x) + 1 - x),$$

which completes the proof. ■

Remark 5. Considering that continuity of T plays no role in Theorem 4, that T has (as non-decreasing function) at most countably many discontinuities, and that $\mu_A(E \times [0, 1]) = 0$ for every countable set E and $A \in \mathcal{C}$ we may, w.l.o.g., assume that T is left continuous, in which case the infimum in eq. (16) is a minimum.

Corollary 6. Suppose that X, Y are random variables with continuous distribution functions F and G respectively, set $T = G \circ F^-$ and $z := 1 + \inf_{x \in [0, 1]} (T(x) - x)$, define $R : [0, 1] \rightarrow [0, 1]$ by $R(x) = z + x \pmod{1}$, and let A_R denote the completely dependent copula induced by R . Then for $(X, Y) \sim H \in \mathcal{H}(F, G)$ with $H(x, y) = A_R(F(x), G(y))$ we have $\mathbb{P}(Y \leq X) = \overline{m}$.

Example 7. Suppose that the default times X and Y are exponentially distributed with parameters θ_1 and θ_2 respectively. It is straightforward to verify that in this case $T = G \circ F^-$ is given by $T_\theta(x) = 1 - (1 - x)^\theta$, where $\theta = \frac{\theta_2}{\theta_1}$. For the case of $\theta \geq 1$ we have $T_\theta(x) \geq x$ for every $x \in [0, 1]$, so $\sup_{A \in \mathcal{C}} \mu_A(\Gamma^\leq(T_\theta)) = 1$. Remarkably, for the case of $\theta < 1$ the maximal mass of the endograph of T_θ and the maximal mass of the graph of T_θ coincide. In fact, applying Theorem 4, on the one hand we get

$$\sup_{A \in \mathcal{C}} \mu_A(\Gamma^\leq(T_\theta)) = 1 + \theta^{\frac{1}{1-\theta}} - \theta^{\frac{\theta}{1-\theta}}.$$

And on the other hand, according to Theorem 3 and Theorem 4 in [1] (also see [11, 16]) we have

$$\sup_{A \in \mathcal{C}} \mu_A(\Gamma(T_\theta)) = \int_{[0, 1]} \left(\mathbf{1}_{[0, 1]}(f \circ T_\theta) + \frac{1}{f \circ T_\theta} \mathbf{1}_{(1, \infty)}(f \circ T_\theta) \right) d\lambda \quad (17)$$

where f denotes the density of λ^{T_θ} . Since for $T_\theta(x)$ the latter is given by $f(x) = \frac{1}{\theta} (1 - x)^{\frac{1-\theta}{\theta}}$, we get $f \circ T_\theta(x) = \frac{1}{\theta} (1 - x)^{1-\theta}$ and eq. (17) calculates to

$$\begin{aligned} \sup_{A \in \mathcal{C}} \mu_A(\Gamma(T_\theta)) &= \int_{[0, 1-\theta^{\frac{1}{1-\theta}}]} \frac{1}{\frac{1}{\theta} (1 - x)^{1-\theta}} d\lambda(x) + 1 - (1 - \theta^{\frac{1}{1-\theta}}) = 1 - \theta^{\frac{\theta}{1-\theta}} + \theta^{\frac{1}{1-\theta}}. \\ &= \sup_{A \in \mathcal{C}} \mu_A(\Gamma^\leq(T_\theta)). \end{aligned}$$

For the special case of $\theta = \frac{1}{2}$, which is depicted in Figure 1, we get

$$\sup_{A \in \mathcal{C}} \mu_A(\Gamma(T)) = \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)) = \frac{3}{4}.$$

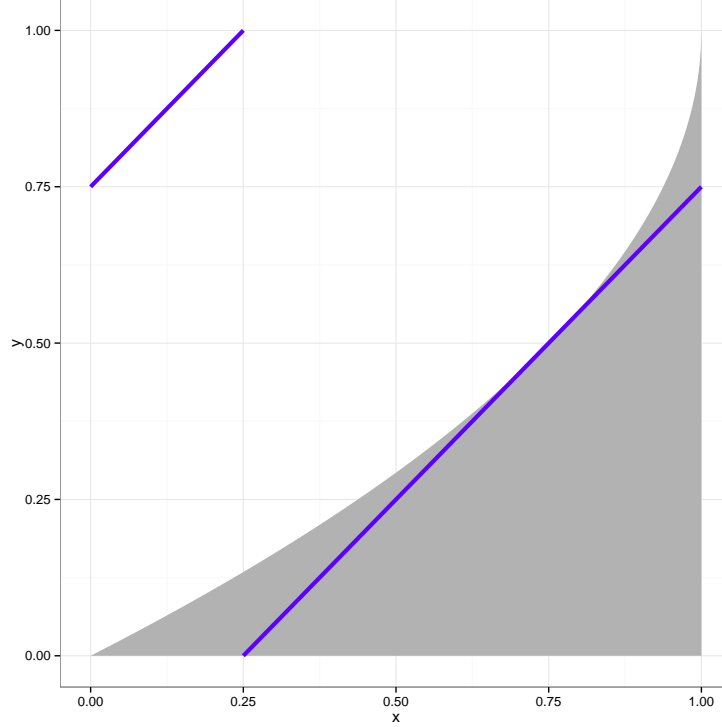


Figure 1: The endograph $\Gamma^{\leq}(T)$ of the transformation $T(x) = 1 - (1 - x)^{\frac{1}{2}}$ (shaded region) and the support of the mutually completely dependent copula A_R constructed in the proof of Theorem 4 assigning maximum mass to $\Gamma^{\leq}(T)$ (blue).

Example 8. Based on Example 7 it might seem natural to conjecture that the equality $\sup_{A \in \mathcal{C}} \mu_A(\Gamma(T)) = \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T))$ holds for a much bigger class of non-decreasing transformations T fulfilling $T(x) \leq x$ for every $x \in [0, 1]$. Since counterexamples are easily constructed for the case where T is singular ($\lambda^T(E) > 0$ for some $E \in \mathcal{B}([0, 1])$ with $\lambda(E) = 0$) and the case where T has discontinuities, the conjecture reduces to strictly increasing, continuous transformations T . For every $n \in \mathbb{N}$ the transformation $T_n : [0, 1] \rightarrow [0, 1]$, defined by

$$T_n(x) = \begin{cases} \frac{x}{2} & \text{if } x \in [0, \frac{1}{2}] \\ \frac{x}{2} + \frac{x}{2} \sqrt[n]{4x - 2} & \text{if } x \in (\frac{1}{2}, \frac{3}{4}) \\ -1 + 2x & \text{if } x \in [\frac{3}{4}, 1] \end{cases}$$

is easily verified to be homeomorphism with $T_n(x) \leq x$ for every $x \in [0, 1]$ (see Figure 2 for the case $n = 10$). Applying Theorem 4 we get $\sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)) = \frac{3}{4}$, however, either by graphical arguments or by using Theorem 3 and Theorem 4 in [1] it is straightforward to verify that $\lim_{n \rightarrow \infty} \sup_{A \in \mathcal{C}} \mu_A(\Gamma(T_n)) = \frac{1}{2} < \frac{3}{4}$, so the conjecture is wrong.

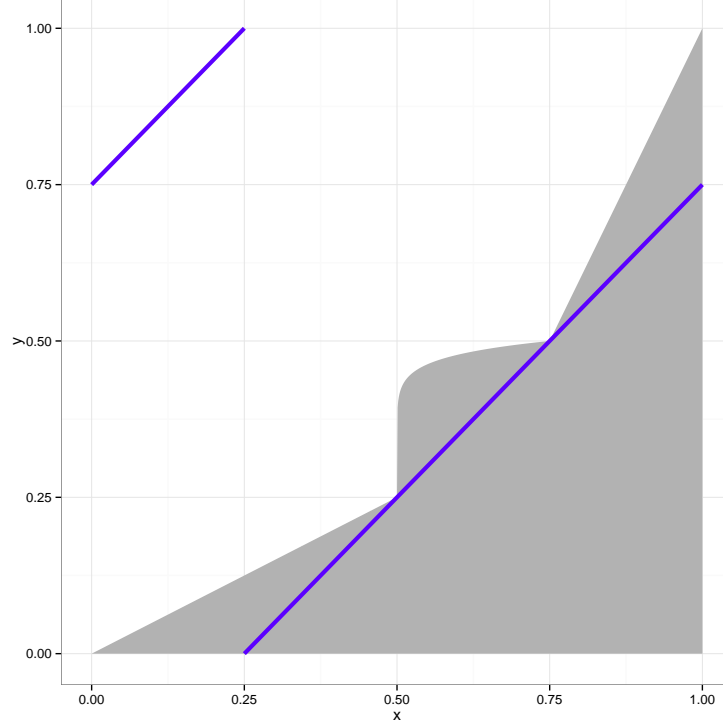


Figure 2: The endograph $\Gamma^{\leq}(T_{10})$ of the transformation T_{10} from Example 8 (shaded region) and the support of the mutually completely dependent copula A_R constructed in the proof of Theorem 4 assigning maximum mass to $\Gamma^{\leq}(T_{10})$ (blue).

Although monotonicity is crucial in the proof of Theorem 4 it is even possible to calculate

$$\overline{m} := \sup_{\mu \in \mathcal{P}(F, G)} \mu(\Gamma^{\leq}(S)) = \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)) \quad (18)$$

for the case of arbitrary measurable (not necessarily monotonic) transformations $S : \mathbb{R} \rightarrow \mathbb{R}$. Letting $T : [0, 1] \rightarrow [0, 1]$ denote an arbitrary measurable transformation, we will now directly concentrate on the quantity \overline{m}_T , defined by

$$\overline{m}_T := \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)) \quad (19)$$

and prove a simple formula for \overline{m}_T only involving the d.f. $F_T : [0, 1] \rightarrow [0, 1]$ of T , defined by

$$F_T(x) = \lambda^T([0, x]) = \lambda(T^{-1}([0, x])). \quad (20)$$

We start with two simple lemmata that will be used in the proof of the main results - the first one contains an alternative simple formula for \overline{m}_T involving F_T which will be key in the proofs of the main results, the second one gathers two properties describing how much \overline{m}_T may change if T changes.

Lemma 9. *Suppose that $T : [0, 1] \rightarrow [0, 1]$ is measurable. Then we have*

$$\overline{m}_T \leq 1 + \inf_{x \in [0, 1]} (T(x) - F_T \circ T(x)) = 1 + \inf_{y \in [0, 1]} (y - F_T(y)) = 1 + \min_{y \in [0, 1]} (y - F_T(y)) \quad (21)$$

If T is non-decreasing then we have equality in (21).

Proof: Considering $\Gamma^{\leq}(T) \subseteq [0, 1] \times [0, T(x)] \cup T^{-1}((T(x), 1]) \times [0, 1]$ and using $\lambda^T((T(x), 1]) = 1 - F_T \circ T(x)$ we get

$$\mu_A(\Gamma^{\leq}(T)) \leq T(x) + 1 - F_T \circ T(x)$$

holds for every $x \in [0, 1]$ and every $A \in \mathcal{C}$, from which the first inequality follows immediately. To prove the second part of (21) it suffices to show that for every $y \in [0, 1]$ we have

$$\inf_{x \in [0, 1]} (T(x) - F_T \circ T(x)) \leq y - F_T(y). \quad (22)$$

In the following $Rg(T)$ will denote the range of T , $\overline{Rg(T)}$ its topological closure. It is easy to see that the left hand-side of ineq. (22) can not exceed zero: In fact, setting $u := \sup(Rg(T))$ there are two possibilities: If $u = T(x)$ for some $x \in [0, 1]$ then $T(x) - F_T(T(x)) = u - 1 \leq 0$. And if $u \notin Rg(T)$ then $\lambda^T(\{u\}) = 0$, so u is a continuity point of F_T and, by construction, we can find a monotonically increasing sequence $(T(x_n))_{n \in \mathbb{N}}$ converging to u , implying $0 \geq u - 1 = u - F_T(u) = \lim_{n \rightarrow \infty} T(x_n) - F_T \circ T(x_n)$.

For $y \in Rg(T)$ and, using the previous paragraph, for $y = 1$ and for $F_T(y) = 0$ ineq. (22) is trivial. The inequality is also clear for $y = 0$ since in case of $F_T(0) > 0$ we have $y \in Rg(T)$. Suppose now that $y \in (0, 1)$ and that $y \notin Rg(T)$. Then obviously $F_T(y) = F_T(y-)$, i.e. y is a continuity point of F_T . Consequently, if $y \in \overline{Rg(T)}$ then there exists a sequence $(T(x_n))_{n \in \mathbb{N}}$ converging to y , so $y - F_T(y) = \lim_{n \rightarrow \infty} T(x_n) - F_T \circ T(x_n) \geq \inf_{x \in [0, 1]} (T(x) - F_T \circ T(x))$. Considering that $y < \inf_{x \in [0, 1]} T(x)$ implies $F_T(y) = 0$, whence ineq. (22), it remains to prove the inequality for the case that $y \notin \overline{Rg(T)}$, $y > \inf_{x \in [0, 1]} T(x)$ and $F_T(y) > 0$. Setting $y_0 = F_T^-(F_T(y))$ we have $y_0 < y$ as well as $F_T(y_0) = F_T(y)$, so $y_0 - F_T(y_0) < y - F_T(y)$. Since the construction of y_0 implies $y_0 \in \overline{Rg(T)}$ the proof of ineq. (22) is complete.

Proving the existence of $y^* \in [0, 1]$ fulfilling $I := \inf_{y \in [0, 1]} (y - F_T(y)) = y^* - F_T(y^*)$ can be done as follows: For every $n \in \mathbb{N}$ we can find $y_n \in [0, 1]$ with $y_n - F_T(y_n) < I + \frac{1}{2n}$. Compactness of $[0, 1]$ implies the existence of a subsequence $(y_{n_j})_{j \in \mathbb{N}}$ and some $y^* \in [0, 1]$ with $\lim_{j \rightarrow \infty} y_{n_j} = y^*$. If $y^* = 1$ we are done since $I = \lim_{j \rightarrow \infty} (y_{n_j} - F_T(y_{n_j})) = y^* - \lim_{j \rightarrow \infty} F_T(y_{n_j}) \geq y^* - 1 = y^* - F_T(y^*)$. Suppose therefore that $y^* < 1$ and let $\delta \in (0, 1 - y^*]$ be arbitrary. Then there exists an index $j_0 \in \mathbb{N}$ such that $y_{n_j} < y^* + \delta$, hence $y_{n_j} - F_T(y_{n_j}) \geq y_{n_j} - F_T(y^* + \delta)$, holds for all $j \geq j_0$. Considering $j \rightarrow \infty$ yields $I \geq y^* - F_T(y^* + \delta)$, hence, using right-continuity of F_T we get $I \geq y^* - F_T(y^*)$.

Finally, suppose that T is non-decreasing. We want to show that

$$\inf_{x \in [0, 1]} (T(x) - F_T \circ T(x)) = \inf_{x \in [0, 1]} (T(x) - x) \quad (23)$$

It follows directly from the construction that $F_T \circ T(x) \geq x$ holds for every $x \in [0, 1]$, so the left-hand side of (23) can not be greater than the right-hand side. Additionally, it is straightforward to verify that $F_T \circ T(x) > x$ holds if and only if there exists $z > x$ with $T(x) = T(z)$. Hence in case of $F_T \circ T(x_0) > x_0$, setting $\langle a, b \rangle = T^{-1}(\{T(x_0)\})$, $x_0 < b$ follows and, using $\lim_{x \rightarrow b-} (T(x) - x) = T(x_0) - b = T(x_0) - F_T \circ T(x_0)$, we get that $\inf_{x \in [0, 1]} (T(x) - x) \leq T(x_0) - F_T \circ T(x_0)$, which completes the proof. ■

Lemma 10. Suppose that $T, T' : [0, 1] \rightarrow [0, 1]$ are measurable transformations. Then the following two assertions hold:

1. For $D := \{x \in [0, 1] : T(x) \neq T'(x)\}$ we have $|\overline{m}_{T'} - \overline{m}_T| \leq \lambda(D)$.
2. If $\Delta \in [0, 1)$ and $T' \geq T - \Delta$, then $\overline{m}_{T'} \geq \overline{m}_T - \Delta$ follows.

Proof: To prove the first assertion set $L := T \mathbf{1}_{D^c}$ and $U := T \mathbf{1}_{D^c} + \mathbf{1}_D$. Considering that obviously

$$\mu_A(\Gamma^\leq(L)) \leq \min \{\mu_A(\Gamma^\leq(T)), \mu_A(\Gamma^\leq(T'))\} \leq \max \{\mu_A(\Gamma^\leq(T)), \mu_A(\Gamma^\leq(T'))\} \leq \mu_A(\Gamma^\leq(U))$$

as well as $0 \leq \mu_A(\Gamma^\leq(U)) - \mu_A(\Gamma^\leq(L)) = \mu_A(D \times [0, 1]) = \lambda(D)$ holds for every $A \in \mathcal{C}$, we get $|\mu_A(\Gamma^\leq(T)) - \mu_A(\Gamma^\leq(U))| \leq \lambda(D)$ for every $A \in \mathcal{C}$. Having this, the desired inequality follows immediately.

To prove the second assertion let $R_\Delta : [0, 1] \rightarrow [0, 1]$ be defined by $R_\Delta(x) = x + \Delta \pmod{1}$ and fix $A \in \mathcal{C}$. Since obviously $R_\Delta \in \mathcal{T}$, defining $\mu(E \times F) = \mu_A(E \times R_\Delta(F))$ yields a doubly stochastic measure μ which corresponds to a copula A_Δ (that, in turn, is easily seen to be the transpose of the R_Δ -shuffle $\mathcal{S}_{R_\Delta}(A)$ of A). Defining $\tilde{T} : [0, 1] \rightarrow [0, 1]$ by $\tilde{T}(x) = \max\{T(x) - \Delta, 0\}$, $\tilde{T} \leq T'$ follows and, using disintegration, we get

$$\begin{aligned} \mu_{A_\Delta}(\Gamma^\leq(T')) &\geq \mu_{A_\Delta}(\Gamma^\leq(\tilde{T})) = \int_{T^{-1}([\Delta, 1])} K_{A_\Delta}(x, [0, T(x) - \Delta]) d\lambda(x) \\ &= \int_{T^{-1}([\Delta, 1])} K_A(x, [\Delta, T(x)]) d\lambda(x) \\ &= \int_{[0, 1]} K_A(x, [0, T(x)]) d\lambda(x) - \int_{T^{-1}([0, \Delta])} K_A(x, [0, T(x)]) d\lambda(x) \\ &\quad - \int_{T^{-1}([\Delta, 1])} K_A(x, [0, \Delta]) d\lambda(x) \\ &\geq \mu_A(\Gamma^\leq(T)) - \int_{[0, 1]} K_A(x, [0, \Delta]) d\lambda(x) = \mu_A(\Gamma^\leq(T)) - \Delta. \end{aligned}$$

Since $A \in \mathcal{C}$ was arbitrary it follows immediately that $\overline{m}_{T'} \geq \overline{m}_T - \Delta$. ■

We now tackle the calculation of \overline{m}_T for arbitrary measurable T in two steps - we first prove the result for continuous T and then extend it via Lusin's theorem (see [14]) and some compactness arguments to the general case. Since the proof for Riemann-integrable T is only slightly more complicated than that for continuous T we directly focus on Riemann-integrable transformations $T : [0, 1] \rightarrow [0, 1]$.

Theorem 11. *Suppose that $T : [0, 1] \rightarrow [0, 1]$ is Riemann-integrable. Then we have*

$$\overline{m}_T = 1 + \min_{x \in [0, 1]} (x - F_T(x)). \quad (24)$$

Proof: Let $w_T(x)$ denote the oscillation of T at the point $x \in [0, 1]$, i.e.

$$w_T(x) = \lim_{r \rightarrow 0^+} \sup_{u, v \in B(x, r)} |T(u) - T(v)|,$$

where $B(x, r) = \{z \in [0, 1] : |z - x| < r\}$. It is well known ([7]) that w_T is upper semicontinuous and that x is a continuity point of T if, and only if $w_T(x) = 0$.

In what follows let $\varepsilon > 0$ be arbitrary but fixed. Riemann-integrability ([7]) of T implies that $E = \{x \in [0, 1] : w_T(x) \geq \varepsilon\}$ is compact and fulfills $\lambda(E) = 0$ so we can find open intervals U_1, \dots, U_n such that $E \subseteq \bigcup_{i=1}^n U_i$ and $\lambda(\bigcup_{i=1}^n U_i) \leq \varepsilon$ holds. Set $K = [0, 1] \setminus \bigcup_{i=1}^n U_i$. For every $x \in K$ we have $w_T(x) < \varepsilon$ and, using compactness of K , we can find pairwise disjoint intervals J_1, \dots, J_m such that $\bigcup_{i=1}^m J_i = K$ and $\sup_{u,v \in J_i} |T(u) - T(v)| < \varepsilon$ holds for every $i \in \{1, \dots, m\}$. Defining a step function $S : [0, 1] \rightarrow [0, 1]$ by

$$S(x) = \begin{cases} \inf_{z \in J_i} T(z) & \text{if } x \in J_i \text{ for some } i \in \{1, \dots, m\}, \\ 0 & \text{otherwise} \end{cases}$$

we have $T(x) - \varepsilon < S(x) \leq T(x)$ for every $x \in K$. Applying Lemma 10 yields $\overline{m}_T \geq \overline{m}_S \geq \overline{m}_T - 2\varepsilon$. Proceeding in completely the same manner we can construct a sequence $(S_n)_{n \in \mathbb{N}}$ of step functions such that $S_n \leq T$ and

$$\overline{m}_T \geq \overline{m}_{S_n} \geq \overline{m}_T - \frac{1}{2^n}$$

holds for every $n \in \mathbb{N}$, which implies $\overline{m}_T = \lim_{n \rightarrow \infty} \overline{m}_{S_n}$. For each n we can reorder the intervals on which S_n is constant in such a way that the resulting step function T_n is monotonically increasing. Working with classical shuffles it is straightforward to verify that $\overline{m}_{S_n} = \overline{m}_{T_n}$ holds. According to Lemma 9 we have $\overline{m}_{T_n} = 1 + \min_{y \in [0,1]} (y - F_{T_n}(y))$, so taking into account $F_{T_n} = F_{S_n}$ altogether we have already shown

$$\lim_{n \rightarrow \infty} \left(1 + \min_{y \in [0,1]} (y - F_{S_n}(y)) \right) = \overline{m}_T,$$

and we are done if we can prove that

$$\lim_{n \rightarrow \infty} \underbrace{\left(\min_{y \in [0,1]} (y - F_{S_n}(y)) \right)}_{=: I_n} = \underbrace{\min_{y \in [0,1]} (y - F_T(y))}_{=: I}. \quad (25)$$

Considering $S_n \leq T$ we have $y - F_{S_n}(y) \leq y - F_T(y)$, so $\lim_{n \rightarrow \infty} I_n \leq I$. The construction of S_n implies $\lim_{n \rightarrow \infty} \|S_n - T\|_1 = 0$, so (see [9, 14]) there exists a subsequence $(S_{n_i})_{i \in \mathbb{N}}$ converging λ -a.e. to T . Using the fact that almost sure convergence implies weak convergence (see [9]) it follows that $\lim_{n \rightarrow \infty} F_{S_{n_i}}(y) = F_T(y)$ holds for every point $y \in [0, 1]$ at which F_T is continuous. Choose y_{n_i} in such a way that $\min_{y \in [0,1]} (y - F_{S_{n_i}}(y)) = y_{n_i} - F_{S_{n_i}}(y_{n_i})$. W.l.o.g. (consider another subsequence if necessary) assume that $(y_{n_i})_{i \in \mathbb{N}}$ converges to some $y^* \in [0, 1]$. For $y^* = 1$ we have $F_T(y^*) = 1$, from which $\lim_{i \rightarrow \infty} I_{n_i} \geq \lim_{i \rightarrow \infty} (y_{n_i} - 1) = y^* - F_T(y^*) \geq I$ follows. Suppose therefore that $y^* < 1$ and let $y \in (y^*, 1)$ denote a continuity point of F_T . Then there exists an index $i_0 \in \mathbb{N}$ such that for all $i \geq i_0$ we have $y_{n_i} < y$, so, in particular, $I_{n_i} \geq y_{n_i} - F_{T_{n_i}}(y)$. Since the latter implies $\lim_{i \rightarrow \infty} I_{n_i} \geq y^* - F_T(y)$, taking into account that the set of all continuity points of F_T is dense, we finally get $\lim_{i \rightarrow \infty} I_{n_i} \geq y^* - F_T(y^*) \geq I$, which completes the proof of the theorem. ■

Theorem 12. *Suppose that $T : [0, 1] \rightarrow [0, 1]$ is measurable. Then we have*

$$\overline{m}_T = 1 + \min_{x \in [0,1]} (x - F_T(x)). \quad (26)$$

Proof: For every $n \in \mathbb{N}$, Lusin's theorem (see [14]) implies the existence of a compact set $E_n \subseteq [0, 1]$ and a continuous (hence Riemann-integrable) function T_n such that $T_n(x) = T(x)$ for all $x \in E_n$ and $\lambda(E_n) > 1 - \frac{1}{2^n}$. Applying Lemma 10 and Theorem 11 immediately yields

$$\overline{m}_T = \lim_{n \rightarrow \infty} \left(1 + \min_{x \in [0,1]} (x - F_{T_n}(x)) \right)$$

and the theorem is proved if we can show that

$$\lim_{n \rightarrow \infty} \underbrace{\left(\min_{x \in [0,1]} (x - F_{T_n}(x)) \right)}_{=: I_n} = \underbrace{\min_{x \in [0,1]} (x - F_T(x))}_{=: I}. \quad (27)$$

For $n \in \mathbb{N}$ and arbitrary $x \in [0, 1]$ we get

$$\begin{aligned} F_{T_n}(x) &= \lambda^{T_n}([0, x]) = \lambda(\{z \in E_n : T_n(z) \leq x\}) + \lambda(\{z \in E_n^c : T_n(z) \leq x\}) \\ &= \lambda(\{z \in [0, 1] : T(z) \leq x\}) - \lambda(\{z \in E_n^c : T(z) \leq x\}) + \lambda(\{z \in E_n^c : T_n(z) \leq x\}) \\ &= F_T(x) + \Delta \end{aligned} \quad (28)$$

for some $\Delta \in (-2^{-n}, 2^{-n})$. Since x was arbitrary this implies that $(F_{T_n})_{n \in \mathbb{N}}$ converges uniformly to F_T , based on which it is straightforward to prove eq. (27): (i) If $I = x^* - F_T(x^*)$ for some $x^* \in [0, 1]$ then eq. (28) and the definition of I_n yield

$$I_n \leq x^* - F_{T_n}(x^*) \leq x^* - F_T(x^*) + \frac{1}{2^n} = I + \frac{1}{2^n},$$

from which $\lim_{n \rightarrow \infty} I_n \leq I$ follows immediately. (ii) To prove the opposite inequality, for every $n \in \mathbb{N}$ choose $x_n \in [0, 1]$ such that $I_n = x_n - F_{T_n}(x_n)$. Applying eq. (28) yields

$$I - \frac{1}{2^n} \leq x_n - F_T(x_n) - \frac{1}{2^n} \leq x_n - F_{T_n}(x_n) = I_n,$$

from which $I \leq \lim_{n \rightarrow \infty} I_n$ follows. ■

5. An alternative proof of the main result and some consequences

Theorem 12 can be proved in a different way by using Lemma 9 and results from [15]. In fact, slightly modifying the ideas in the first Section of [15] it can be shown that for each measurable $T : [0, 1] \rightarrow [0, 1]$ there exists a non-decreasing function $T^* : [0, 1] \rightarrow [0, 1]$ (called the non-decreasing rearrangement of T) and a λ -preserving transformation $\varphi : [0, 1] \rightarrow [0, 1]$ such that

$$T^* \circ \varphi = T \quad (29)$$

holds. Having this, letting $\mathcal{U}_\varphi : \mathcal{C} \rightarrow \mathcal{C}$ denote the operator studied in [18] and implicitly defined via

$$K_{\mathcal{U}_\varphi(A)}(x, E) = K_A(\varphi(x), E),$$

and using disintegration and change of coordinates we get that

$$\begin{aligned} \mu_{\mathcal{U}_\varphi(A)}(\Gamma^\leq(T)) &= \int_{[0,1]} K_{\mathcal{U}_\varphi(A)}(x, [0, T(x)]) d\lambda(x) = \int_{[0,1]} K_A(\varphi(x), [0, T^* \circ \varphi(x)]) d\lambda(x) \\ &= \int_{[0,1]} K_A(z, [0, T^*(z)]) d\lambda(z) = \mu_A(\Gamma^\leq(T^*)) \end{aligned} \quad (30)$$

holds for every $A \in \mathcal{C}$, implying $\overline{m}_T \geq \overline{m}_{T^*}$. Again using $T^* \circ \varphi = T$ and the fact that φ is λ -preserving, it is straightforward to verify that T and T^* have the same d.f., i.e. $F_{T^*} = F_T$ holds. Therefore, applying Lemma 9 yields

$$1 + \min_{x \in [0,1]} (x - F_T(x)) = 1 + \min_{x \in [0,1]} (x - F_{T^*}(x)) = \overline{m}_{T^*} \leq \overline{m}_T \leq 1 + \min_{x \in [0,1]} (x - F_T(x)), \quad (31)$$

from which the desired equality $\overline{m}_{T^*} = \overline{m}_T$ follows immediately. Although this alternative proof is shorter we opted for the one presented in the previous section since, firstly, it is self-contained and, secondly, Lemma 10 is interesting in itself and will also be used in the sequel when deriving some corollaries.

According to Theorem 4 the completely dependent copula $A_R \in \mathcal{C}_d$ fulfills $\overline{m}_{T^*} = \mu_{A_R}(\Gamma^{\leq}(T^*))$, so eq. (30) implies $\mu_{\mathcal{U}_{\varphi}(A_R)}(\Gamma^{\leq}(T)) = \mu_{A_R}(\Gamma^{\leq}(T^*)) = \overline{m}_{T^*} = \overline{m}_T$. By definition of $\mathcal{U}_{\varphi}(C)$ we have

$$K_{\mathcal{U}_{\varphi}(A_R)}(x, F) = K_{A_R}(\varphi(x), F) = \mathbf{1}_F(R \circ \varphi(x)), \quad (32)$$

so $\mathcal{U}_{\varphi}(A_R)$ is completely dependent and the following corollary holds:

Corollary 13. *Suppose that $T : [0, 1] \rightarrow [0, 1]$ is measurable. Then there exists a completely dependent copula $A_h \in \mathcal{C}_d$ such that $\mu_{A_h}(\Gamma^{\leq}(T)) = \overline{m}_T$.*

Having found a simple and easily computable formula for the maximal mass of $\Gamma^{\leq}(T)$ we now derive the analogous result for the minimal mass and set

$$\underline{m}_T = \inf_{A \in \mathcal{C}} \mu_A(\Gamma^{\leq}(T)). \quad (33)$$

Given the aforementioned results, the subsequent corollary does not come as a surprise:

Corollary 14. *For every measurable transformation $T : [0, 1] \rightarrow [0, 1]$ the following equality holds:*

$$\underline{m}_T = \max_{x \in [0,1]} (x - F_T(x)) \quad (34)$$

Proof: We first concentrate on the *strict endograph* $\Gamma^{<}(T)$, defined by

$$\Gamma^{<}(T) = \{(x, y) \in [0, 1]^2 : y < T(x)\}.$$

Defining $T_n : [0, 1] \rightarrow [0, 1]$ by $T_n(x) = \max\{T(x) - 2^{-n}, 0\}$ for every $x \in [0, 1]$ and $n \in \mathbb{N}$ we obviously have that $(\Gamma^{\leq}(T_n))_{n \in \mathbb{N}}$ is monotonically increasing and that $\Gamma^{<}(T) = \bigcup_{n=1}^{\infty} \Gamma^{\leq}(T_n)$. Lemma 10 yields $\overline{m}_{T_n} \geq \overline{m}_T - 2^{-n}$ and Corollary 13 implies the existence of a copula $A_n \in \mathcal{C}_d$ with $\mu_{A_n}(\Gamma^{\leq}(T_n)) = \overline{m}_{T_n}$. Altogether we get

$$\overline{m}_{T_n} = \mu_{A_n}(\Gamma^{\leq}(T_n)) \leq \mu_{A_n}(\Gamma^{<}(T)) \leq \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{<}(T)) \leq \overline{m}_T,$$

so considering $n \rightarrow \infty$ shows that $\sup_{A \in \mathcal{C}} \mu_A(\Gamma^{<}(T)) = \overline{m}_T$. Having this, eq. (34) is a straightforward consequence since

$$\underline{m}_T = 1 - \sup_{A \in \mathcal{C}} \mu_A(\Gamma^{<}(1 - T)) = 1 - \overline{m}_{1-T} = - \min_{x \in [0,1]} (x - F_{1-T}(x)) = \max_{x \in [0,1]} (x - F_T(x)). \blacksquare$$

We close the paper with two examples - the first one shows that \underline{m}_T is not necessarily attained whereas the second one considers a non-monotonic transformation for which copulas attaining \underline{m}_T and \overline{m}_T can easily be constructed.

Example 15. For $T(x) = x$, considering rotations R_Δ and the corresponding shuffles $\mathcal{S}_{R_\Delta}(M)$, it follows immediately that $\underline{m}_T = 0$. There is, however, no copula A fulfilling $\mu_A(\Gamma^\leq(T)) = 0$, i.e. contrary to \overline{m}_T , there are situations, in which \underline{m}_T is not attained for any copula. Suppose, on the contrary, that $A \in \mathcal{C}$ fulfills $\mu_A(\Gamma^\leq(T)) = 0$. Then, defining $h \in \mathcal{T}_b$ by $h(x) = 1 - x$ and setting $B = \mathcal{U}_h(A)$, we have $\mu_B(\Gamma^\leq(1 - T)) = 0$, so, $B(x, 1 - x) = 0$ holds for every $x \in [0, 1]$. The latter implies $B = W$, which is a contradiction.

Example 16. For $T(x) = 4(x - \frac{1}{2})^2$ it is straightforward to find mappings T^* and φ such that eq. (29) holds. In fact, defining $\varphi : [0, 1] \rightarrow [0, 1]$ by

$$\varphi(x) = \begin{cases} 1 - 2x & \text{if } x \in [0, \frac{1}{2}] \\ -1 + 2x & \text{if } x \in [\frac{1}{2}, 1] \end{cases}$$

and setting $T^*(x) = x^2$ we immediately get $T^* \circ \varphi = T$. Using eq. (32), and setting $R(x) = x + \frac{3}{4} \pmod{1}$, it follows that $h = R \circ \varphi$ is λ -preserving and that $A_h \in \mathcal{C}_d$ fulfills $\mu_{A_h}(\Gamma^\leq(T)) = \overline{m}_T = \overline{m}_{T^*} = \frac{3}{4}$. Considering that for A_φ we obviously have $\mu_{A_\varphi}(\Gamma^\leq(T)) = 0$, we get $\underline{m}_T = 0$ which coincides with $\max_{x \in [0, 1]} (x - F_T(x))$. Figure 3 depicts the supports of the copulas A_h and A_φ as well as the endograph of T .

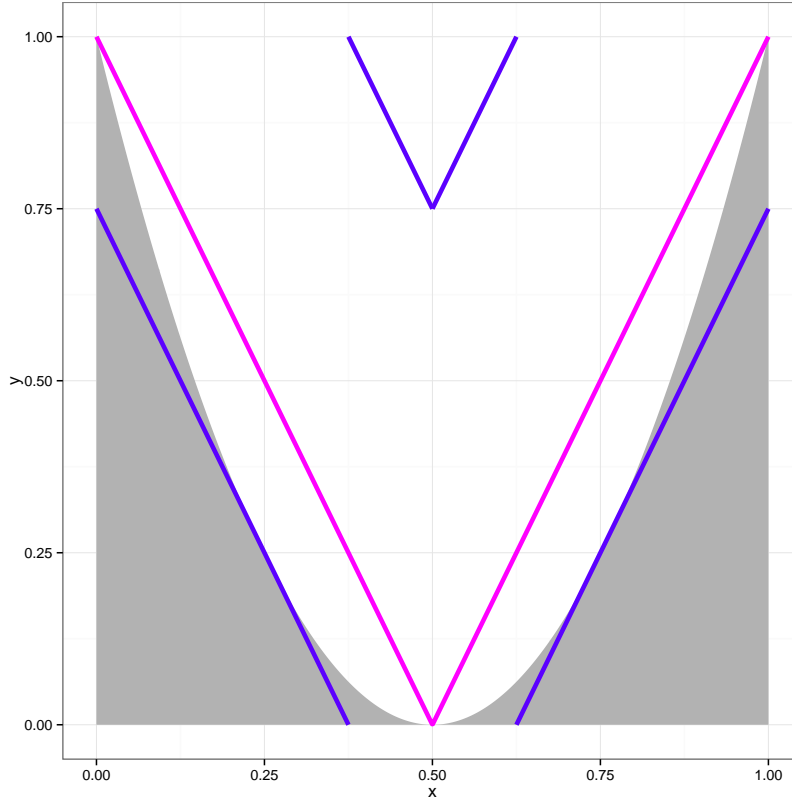


Figure 3: The endograph $\Gamma^\leq(T)$ of the transformation T from Example 16 (shaded region) as well as the support of the copulas A_h and A_φ maximizing/minimizing the mass of $\Gamma^\leq(T)$ (blue and magenta respectively).

Acknowledgments

The third author gratefully acknowledges the support of the grant MTM2014-60594-P (partially supported by FEDER) from the Spanish Ministry of Economy and Competitiveness.

References

- [1] F. Durante, J. Fernández Sánchez, W. Trutschnig: On the singular components of a copula, *J. Appl. Probab.* **52**, 1175-1182 (2015).
- [2] F. Durante, P. Sarkoci, C. Sempi: Shuffles of copulas, *J. Math. Anal. Appl.* **352**, 914-921 (2009).
- [3] F. Durante, C. Sempi: *Principles of Copula Theory*, Chapman and Hall/CRC, 2015.
- [4] P. Embrechts, G. Puccetti: Bounds for functions of dependent risks, *Finance and Stoch* **10**, 341-352 (2006).
- [5] P. Embrechts, G. Puccetti: Bounds for functions of multivariate risks, *J. Multivariate Anal.* **97**, 526-547 (2006).
- [6] P. Embrechts, M. Hofert: A note on generalized inverses, *Math. Method. Oper. Res.* **77**, 423-432 (2013).
- [7] E. Hewitt, K. Stromberg: *Real and Abstract Analysis*, Springer Verlag, Berlin Heidelberg, 1965
- [8] O. Kallenberg: *Foundations of modern probability*, Springer Verlag, New York Berlin Heidelberg, 1997.
- [9] A. Klenke: *Probability Theory - A Comprehensive Course*, Springer Verlag, Berlin Heidelberg, 2007.
- [10] H.O. Lancaster: Correlation and complete dependence of random variables, *Ann. Math. Stat.* **34**, 1315-1321 (1963).
- [11] J.F. Mai, M. Scherer: Simulating from the copula that generates the maximal probability for a joint default under given (inhomogeneous) marginals, in *Topics from the 7th International Workshop on Statistical Simulation* ed. V. Melas, S. Mignani, P. Monari, and L. Salmaso, Springer Proceedings in Mathematics & Statistics **114**, pp. 333-341, 2014.
- [12] P. Mikusiński, H. Sherwood, M.D. Taylor: Shuffles of Min, *Stochastica* **290**, 61-74 (1992).
- [13] R.B. Nelsen: *An Introduction to Copulas*, Springer, New York, 2006.
- [14] W. Rudin: *Real and Complex Analysis*, McGraw-Hill International Editions, Singapore, 1987.
- [15] J.V. Ryff: Measure Preserving Transformations and Rearrangements, *J. Math. Anal. Appl.* **31**, 449-458 (1970).

- [16] H. Thorisson: *Coupling, stationarity, and regeneration*, Probability and its Applications, Springer-Verlag, New York, 2000.
- [17] W. Trutschnig: On a strong metric on the space of copulas and its induced dependence measure, *J. Math. Anal. Appl.* **384**, 690-705 (2011)
- [18] W. Trutschnig, J. Fernández Sánchez: Some results on shuffles of two-dimensional copulas, *J. Stat. Plan. Infer.* **143**, 251-260 (2013).